Appendix B: Water Quality Data Methodology

This appendix describes the methods used to compile the database used to assess water quality in the watershed. Historical as well as current water quality data was obtained from several sources and then analyzed according to the methodology shown in Figure B.1.

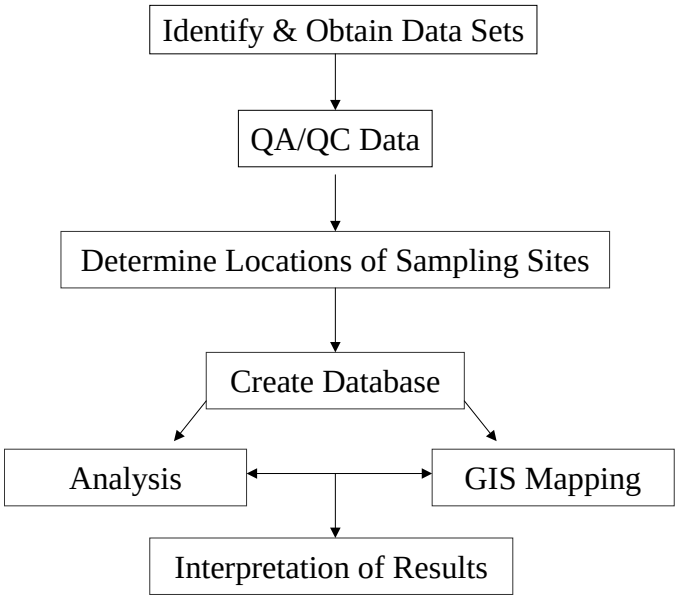

Figure B.1. Water quality data analysis flow chart

Each of the steps in the methodology is described below.

Step 1 – Compile Data

Water quality data collected between 1967 (the earliest year of record) and 2002 were obtained from the written reports and electronic databases (Excel files) are listed in Table B.1

Table B.1. Sources of water quality data (1967-2002)

Written Reports	QA/QA Measures
MWRC, 1970.	*followed Standards Methods for the Examination of Water and Wastewater (12th edition, 1965) *data tabulated and verified by engineers of the Division of Water Pollution Control
MWRC, 1974.	*followed Standards Methods for the Examination of Water and Wastewater (13th edition, 1971) *data tabulated and verified by engineers of the Division of Water Pollution Control
MDC, 1977.	*followed Standards Methods for the Examination of Water and Wastewater (13th edition)
MDC, 1978.	*followed Standards Methods for the Examination of Water and Wastewater (13th edition)

MDC, 1979.	*followed Standards Methods for the Examination of Water and Wastewater (14th edition, 1975, New York)
MDC, 1980.	
MDC, 1981.	
MDC, 1982.	
DEQE, 1982	*followed Standards Methods for the Examination of Water and Wastewater (14th edition, 1975, New York)
MADEP, 1989(1)	*followed Standards Methods for the Examination of Water and Wastewater (16th edition, 1985, New York)
MADEP, 1989(2)	*followed Standard Operating Procedures, Basin Planning Section 1988, Standard Methods for the Examination of Water and Wastewater
.	
Electronic Databases	
MWRA, 1989-2002[1]	MADEP-approved QAPP
MMN, 2000-2002[2]	MADEP-approved QAPP
USGS-NAWQA, 1998-2002[3]	
USGS, 1999-2000[4]	
Tufts, 2000[5]	USEPA-approved QAPP
Tufts, 2002[6]	USEPA-approved QAPP

[1]Datafile received by personal communication from Kelly Coughlin, MWRA, Charlestown, MA, 2002.
[2]Mystic Monitoring Network (MMN), datafile received by personal communication from Libby Larson, MyRWA, Arlington, MA 2003.
[3]Data was downloaded from the USGS NAWQA website (http://waterdata.usgs.gov/nwis/qwdata)
[4]Datafile received by personal communication from Leslie DeSimone, USGS, Marlborough, MA, 2001.
[5]Datafile received by personal communication from Elizabeth Higgins, Tufts University, Medford, MA, 2001.
[6]Datafile received by personal communication from Kim Oriel, Tufts University, Medford, MA, 2003.

Step 2 – QA/QC Data

The QA/QC measures used in each investigation for collecting and analyzing water samples are also listed in Table B.1. Data in the written reports was generated before 1988, and the methods used for QA/QC are not well described. The more recent data from MWRA, USGS, Mystic Monitoring Network, and Tufts University investigators was collected by following MADEP- or USEPA-approved Quality Assurance Project Plans, and therefore, we believe that this data is more reliable than the pre-1988 data.

Step 3 – Determine Locations of Sampling Sites

In the 11 written reports the authors used a river-mile system to identify each location. However, they did not specify the starting point of the river-mile system in the reports, and therefore it was difficult to identify the exact location of the sites. Fortunately, the reports contained written descriptions of the sampling locations (e.g., cross streets or identifying markers and the city in which the site resides), and these were used to create a map in GIS of the sites. Some of the electronic files (e.g., those from MWRA and MMN) contained the latitude and longitude of the sampling sites. The latitude and longitude of the USGS sites were found on the USGS website. The location of the Tufts sampling sites was based on physical descriptions contained in the datasets.

According the data sources listed in Table B.1, a total of 88 sites in the watershed have been monitored at least once between 1967 and 2002. A description of the sampling sites and studies that were included for each site are provided by subbasin in Appendix C.

To facilitate mapping the sites, we gave each one a unique address based on the name of the waterbody and the distance in river-kilometers from a specified zero point. Figures C.1 through C.9 in Appendix C show the sampling locations by subbasin.

Step 4 – Create Database

Because the majority of the sites have been sampled several times over the years and as many as a dozen parameters were measured in each sample, a database was setup to facilitate data analysis. Setting up the database involved several steps. First, each report and dataset was studied to determine which parameters had been analyzed. Next, the common reporting unit for each parameter was identified. In most cases this corresponded to the units of milligrams per liter; however, many of the files reported concentrations in terms of moles per liter, and thus conversion was necessary. The details of the units conversion step are given in Table B.3. [to be provided] For values that were at or near the detection limit that contained qualifiers (< or >), we created a separate column in the database for the qualifiers and used the detection limit as the actual value of the measured parameter. For example, if total phosphorus was reported as "<0.05 mg/L", we assumed that it was present at 0.05 mg/L. The final step in setting up the database was to assign a unique sampling site label to each record (described in Step 3) and then merging all of the datasets into a single flatfile. Using this flatfile and the sampling location information in Table B.2, a database was created in Access®, a database software package.

Steps 5 & 6 – Data analysis/mapping in GIS

Analysis of the data from each subbasin was done by performing queries for violations of the water quality standards and guidelines. Analysis of temporal and spatial trends was performed by querying the data for specific sites and parameters. The results of the queries were then tabulated and in some cases graphed and mapped in ArcMap® thereby allowing more in depth analysis of the data. These results are reported in Chapter 4 and Appendix C.